

User trust and credibility of conversational agents

Jieyu Wang
*Mark B. Schmidt**
Raunak Kuikel
St. Cloud State University
mark@stcloudstate.edu

Extended Abstract

Conversational agents (CA) are the natural language conversational systems that can interact with users intelligently. They are gaining popularity because of the potential benefits for users, such as information retrieval and decision-making. There is limited research that studies users' trust and the credibility of current conversational agents. Our study aims to explore to what extent users trust or do not trust conversational agents and the reasons that explain these behaviors together with CAs' credibility from users' perspectives. Our study provides users' feedback and sheds light on the future finer design for CAs.

Our study explores users' trust and credibility towards CAs. To address the research goal above, we invited 29 participants who are college students to chat with three CAs, Kuki, Replika, and Cleverbot. These three chatbots were rated as the top three ones that these participants prefer to interact with. The participants were assigned the same two tasks for each of the three chatbots. One task is about travel planning and the other one concerns ordering food from restaurants.

Task 1: The spring break is coming. You are pretty interested in traveling. But you do not know where to travel. Please talk to each CA: 1)Kuki, 2)Replika, and 3) Cleverbot and gather enough information for you to create your travel itinerary (a detailed travel plan).

Task 2: Today you are too tired to cook. Also, you would like to explore restaurants. Talk to the three CAs and get your food.

The participants were required to record their conversation history and rate each CA's response to their questions or interactions on a Word document within two weeks. We asked the participants to use a Likert scale of 7 (1=strongly untrustworthy/trustful?, 2=untrustworth, 3=moderately untrustworthy, 4=undecided, 5=moderately trustworthy, 6=trustworthy, 7=strongly trustworthy) to rate each CA response. The participants were also required to provide reasons (written in text) for each response. This large amount of data allows us to explore the participants' extent of trust towards different chatbots' responses and the reasons explaining their behaviors. Currently, we are categorizing the participants' reasons for rating, such as trust, credibility, privacy concerns, etc. We later will use quantitative and qualitative data analysis methods to analyze the ratings and the content.

Our work contributes to the fast development of CA design with detailed user ratings and rich reasons to explain their behaviors. We aim to help CA designers understand users' trust and credibility of CAs for better user-centered design.